The Lexicographic First Occurrence of a I-II-III-Pattern

Torey Burton, Anant P. Godbole and Brett M. Kindle
Department of Mathematics
East Tennessee State University*

February 2, 2008

Abstract

Consider a random permutation $\pi \in \mathcal{S}_n$. In this paper, perhaps best classified as a contribution to discrete probability distribution theory, we study the *first* occurrence $X = X_n$ of a I-II-III-pattern, where "first" is interpreted in the lexicographic order induced by the 3-subsets of $[n] = \{1, 2, \ldots, n\}$. Of course if the permutation is I-II-III-avoiding then the first I-II-III-pattern never occurs, and thus $\mathbb{E}(X) = \infty$ for each n; to avoid this case, we also study the first occurrence of a I-II-III-pattern given a bijection $f : \mathbf{Z}^+ \to \mathbf{Z}^+$.

1 Introduction

Consider a random permutation $\pi \in \mathcal{S}_n$. In this short note, perhaps best classified as a contribution to discrete probability distribution theory (AMS Subject Classification 60C05), we study the *first* occurrence $X = X_n$ of a I-II-III-pattern, defined as follows: Order the 3-subsets of $[n] = \{1, 2, \ldots, n\}$ in the "obvious" lexicographic fashion

$$\{1,2,3\} < \{1,2,4\} < \{1,2,5\} < \ldots \{1,2,n\} < \{1,3,4\} < \{1,3,5\} < \ldots$$
$$< \{1, n-1, n\} < \{2,3,4\} < \ldots \{n-2, n-1, n\}.$$

*E-Mail for Corresponding Author: godbolea@etsu.edu

We say that the first I-II-III-pattern occurs at $\{a,b,c\}$ if $\pi(a) < \pi(b) < \pi(c)$ and if $\pi(d) < \pi(e) < \pi(f)$ does not hold for any $\{d,e,f\} < \{a,b,c\}$. Of course if the permutation is I-II-III-avoiding, which occurs ([1]) with probability $\binom{2n}{n}/(n+1)!$, then $X = \infty$ and the first I-II-III-pattern never occurs. Consequently $\mathbb{E}(X) = \infty$ for each n; to avoid this case, we also simultaneously present results on the first occurrence of a I-II-III-pattern given a bijection $f : \mathbf{Z}^+ \to \mathbf{Z}^+$.

2 Results

In what follows, we use the notation $X = abc$ as short for the event $\{X = \{a,b,c\}\}$, and refer to the case of a bijection on $\mathbf{Z}^+$ as the $n = \infty$ case.

Proposition 1 *For each $n \leq \infty$,*

$$\mathbb{P}(X = 12r) = \frac{1}{r-1} - \frac{1}{r}.$$

Proof. Let $\pi(1), \pi(2), \ldots \pi(r)$ be ordered increasingly as $x_1 < x_2 < \ldots x_r$. Then, if $\pi(2) = x_{r-1}$ and $\pi(r) = x_r$, we clearly have $X = 12r$. Conversely if the second largest of $\{\pi(1), \pi(2), \ldots \pi(r)\}$ is not in the 2$^{\text{nd}}$ spot, then either $\pi(2) = x_r$, or $\pi(2) < x_{r-1}$. In the former case, the only way that we can have $X = 12s$ is with $s > r$. If $\pi(2) < x_{r-1}$, there are two possibilities: If $\pi(1) < \pi(2)$, then $X = 12s$ for some $s < r$, and if $\pi(1) > \pi(2)$ then $X \neq 12s$ for any s. Thus we must have $\pi(2) = x_{r-1}$. Now for X to equal $12r$, we must necessarily have $\pi(r) = x_r$, or else we would have an earlier occurrence of a I-II-III pattern. It now follows that $X = 12r$ if and only if $\pi(2) = x_{r-1}$ and $\pi(r) = x_r$ with the other values arbitrary, so that $\mathbb{P}(X = 12r) = \frac{(r-2)!}{r!} = \frac{1}{r-1} - \frac{1}{r}$, as desired. $\square$

Proposition 2 *If $n = \infty$, then $X = 1sr$ for some $2 \leq s < r$ with probability one.*

Proof. Obvious. No matter what value $f(1)$ assumes, there is an s with $f(s) > f(1)$ and an $r > s$ with $f(r) > f(s)$. Let s_0, r_0 be the smallest such indices; this yields $X = 1s_0r_0$. $\square$

Our ultimate goal is to try to determine the entire probability distribution of X; for $n = 6$, for example, we can check that the ensemble $\{\mathbb{P}(X = abc) : 1 \leq a < b < c \leq 6\}$ is as follows:

Table 1

The First Occurrence of a I-II-III Pattern when $n = 6$

First I-II-III Pattern	Probability	Cumulative Probability
123	120/720	0.1666
124	60/720	0.2500
125	36/720	0.3000
126	24/720	0.3333
134	50/720	0.4028
135	28/720	0.4417
136	18/720	0.4667
145	26/720	0.5028
146	16/720	0.5250
156	16/720	0.5472
234	48/720	0.6139
235	22/720	0.6444
236	12/720	0.6611
245	24/720	0.6944
246	12/720	0.7111
256	14/720	0.7306
345	24/720	0.7639
346	10/720	0.7778
356	14/720	0.7972
456	14/720	0.8167
∞, i.e. never	132/720	1.0000

Recall that the *median* of any random variable X is any number m such that $\mathbb{P}(X \leq m) \geq 1/2$ and $\mathbb{P}(X \geq m) \geq 1/2$. Now Propositions 1 and 2 together reveal that for $n = \infty$,

$$\mathbb{P}(X \leq 134) \geq \sum_{r=3}^{\infty} \frac{1}{r-1} - \frac{1}{r} = \frac{1}{2}$$

and

$$\mathbb{P}(X \geq 134) = 1 - \sum_{r=3}^{\infty} \frac{1}{r-1} - \frac{1}{r} = \frac{1}{2},$$

which shows that X has 134 as its unique median. For finite n, however, the median is larger – Table 1 reveals, for example, that $m = 145$ for $n = 6$.

Proposition 3 *For $n = \infty$,*

$$\mathbb{P}(X = 1sr) = \frac{1}{(s-1)(r-1)r} = \frac{1}{s-1}\left(\frac{1}{r-1} - \frac{1}{r}\right).$$

Proof. It can easily be proved, as in Proposition 1 and keeping in mind that $n = \infty$, that $X = 1sr$ iff $\pi(s) = x_{r-1}, \pi(r) = x_r$; and $\pi(1) = \max_{1\le j\le s-1} \pi(j)$. It now follows that

$$\mathbb{P}(X = 1sr) = \frac{\binom{r-2}{s-1}(s-2)!(r-s-1)!}{r!} = \frac{1}{(s-1)(r-1)r},$$

as claimed. $\square$

Proposition 3 provides us with the entire distribution of X when $n = \infty$; note that

$$\sum_{s=2}^{\infty}\sum_{r=s+1}^{\infty} \frac{1}{s-1}\left(\frac{1}{r-1} - \frac{1}{r}\right) = \sum_{s=2}^{\infty} \frac{1}{s-1} - \frac{1}{s} = 1.$$

The probability of the first I-II-III pattern occurring at positions $12r$ is the same for all $n \le \infty$, noting, though, that for finite n, $\sum_s \mathbb{P}(X = 12s) = \frac{1}{2} - \frac{1}{n}$. There is, however, a subtle and fundamental difference in general between $\mathbb{P}(X = 1rs), r \ge 3$, when $n = \infty$ and when n is finite. We illustrate this fact for $\mathbb{P}(X = 13r)$ when $n < \infty$. Recall from the proof of Proposition 3 that for X to equal $13r$ in the infinite case, we had to have $\pi(3) = x_{r-1}, \pi(r) = x_r$, and $\pi(2) < \pi(1)$. The above scenario *will still*, in the finite case, cause the first I-II-III pattern to occur at positions $13r$, but there is another case to consider. If $n = \pi(2) > \pi(1)$ then it is *impossible* for X to equal $12s$ for any s; in this case we must have $\pi(3) = x_{r-2}$ and $\pi(r) = x_{r-1}$. The probability of this second scenario is

$$\frac{(r-3)!}{n(r-1)!} = \frac{1}{n(r-2)(r-1)}.$$

Adding, we see that

$$\mathbb{P}(X = 13r) = \frac{1}{2}\left(\frac{1}{r-1} - \frac{1}{r}\right) + \frac{1}{n}\left(\frac{1}{r-2} - \frac{1}{r-1}\right),$$

and, in contrast to the $n=\infty$ case where the net contribution of $\mathbb{P}(X = 13r; r \geq 4)$ was $1/6$, we have

$$\begin{aligned}\sum_{r=4}^{n} \mathbb{P}(X=13r) &= \frac{1}{2}\sum_{r=4}^{n}\left(\frac{1}{r-1}-\frac{1}{r}\right)+\frac{1}{n}\sum_{r=4}^{n}\left(\frac{1}{r-2}-\frac{1}{r-1}\right)\\ &= \frac{1}{6}-\frac{1}{n(n-1)}.\end{aligned}$$

The above example illustrates a general fact:

Theorem 4 *For finite n,*

$$\mathbb{P}(X=1sr)=\sum_{k=0}^{s-2}\frac{\binom{s-2}{k}\binom{r-k-2}{s-k-1}(s-k-2)!(r-s-1)!}{n(n-1)\dots(n-k+1)(r-k)!}.$$

Proof. We may have k of the quantities $\pi(2),\pi(3),\dots,\pi(s-1)$ being greater than $\pi(1)$, where k ranges from 0 to $s-2$. In this case, these πs must equal, from left to right, $(n,n-1,\dots,n-k+1)$. Arguing as before, we must have $\pi(s)=x_{r-1-k}$ and $\pi(r)=x_{r-k}$. The rest of the proof is elementary. $\square$

Unlike the infinite case, $\sum_s\sum_r \mathbb{P}(X=1sr)\neq 1$. So how much *is* $\mathbb{P}(X \geq 234)$, or alternatively, how close to unity is

$$\mathbb{P}(X=1sr)=\sum_{s\geq 2}\sum_{r\geq s+1}\sum_{k=0}^{s-2}\frac{\binom{s-2}{k}\binom{r-k-2}{s-k-1}(s-k-2)!(r-s-1)!}{n(n-1)\dots(n-k+1)(r-k)!}?$$

We obtain the answer in closed form as follows: Conditioning on $\pi(1)$, we see that $X\geq 234$ iff for $j=n,n-1,\dots,1$, $\pi(1)=j$, and the integers $n,n-1,\dots,j+1$ appear from left to right in π. Summing the corresponding probabilities $1/n$, $1/n$,$1/(2!n)$, $1/(3!n)$, etc yields

Proposition 5

$$\mathbb{P}(X\geq 234)\sim\frac{e}{n}.$$

3 Open Problems

- Lexicographic ordering is not our only option; in fact it is somewhat unnatural. Consider another possibility: What is

$$\inf\{k : \text{there is a I} - \text{II} - \text{III pattern in } (\pi(1),\dots,\pi(k))\}?$$

This question is not too hard to answer from Stanley-Wilf theory. Since

$$\mathbb{P}((\pi(1),\ldots,\pi(k)) \text{ is I} - \text{II} - \text{III free}) = \frac{\binom{2k}{k}}{(k+1)!},$$

the probability that k is the first integer for which $(\pi(1),\ldots,\pi(k))$ contains a I-II-III pattern is

$$\frac{\binom{2k-2}{k-1}}{k!} - \frac{\binom{2k}{k}}{(k+1)!}.$$

A more interesting question is the following: Conditional on the fact that first I-II-III pattern occurs only after the kth "spot" is revealed, what is the distribution of the first 3-subset, interpreted in the sense of this paper, that causes this to occur? For example, if we let $n = 6$, and are told that the first k for which there is a I-II-II pattern in $(\pi(1),\ldots,\pi(k))$ is 5, what is the chance that the first set that causes this to happen is $\{1,2,5\}$, $\{1,3,5\}$, $\{1,4,5\}$, $\{2,3,5\}$, $\{2,4,5\}$ or $\{3,4,5\}$?

- Can the results of this paper be readily generalized to other patterns of length 3? To patterns of length 4?

- Theorem 4 and Proposition 5 fall short of providing the exact probabilities $\mathbb{P}(X = rst)$ for $r \geq 2$. Can these admittedly small probabilities be computed exactly or to a high degree of precision?

- Does the distribution of X consist, as it does for $n = 6$, of a series of decreasing segments with the initial probability of segment $j + 1$ no smaller than the final probability of segment j?

Acknowledgment The research of the second named author was supported by NSF Grants DMS-0139286 and DMS-0552730.

References

[1] M. Bona, *Combinatorics of Permutations*, Chapman and Hall/CRC, Boca Raton, 2004.